



驱动AI的可持续未来

文 桑杰·波德尔、沙拉布·库马尔·辛格、马修·吉朱、张逊

提要

随着AI的广泛应用，其带来的能源消耗、碳排放与水资源需求也在同步高速增长。如果不加以控制，AI的环境足迹可能会威胁企业的可持续发展甚至是长期的价值创造。埃森哲依据可持续AI商数（SAIQ）这一新指标，助力企业实现AI的可持续发展。

如今，AI已不再是一个新兴趋势，它正在重塑企业的方方面面。埃森哲研究显示，企业若能完成生成式AI全业务规模化部署，则有望在18个月内实现生产力提升13%，营收增长12%，客户体验改善11%。¹但这一切的前提是AI部署本身必须是可持续且高效的。

然而，随着AI的广泛应用，其带来的能源消耗、碳排放与水资源需求也在同步高速增长。如果不加以控制，这一趋势不仅会威胁地球环境，还会影响企业的长期价值创造。归根结底，企业AI发展的关注点需从“AI性能有多强”转向“AI的资源投入最终能带来哪些实际回报”。

AI的环境足迹：日益凸显的风险

AI技术的高效运行，离不开底层基础设施支撑。从高耗能的数据中心到配套的冷却系统，这些关键设施不仅消耗着大量电力和水资源，还会产生大量的碳排放。这种日益加重的环境负担，正逐步侵蚀AI技术本应带来的经济价值。

从具体数据来看，AI对资源的消耗规模呈现出惊人的增长态势。有研究测算显示，到2030年，全球AI数据中心的年耗电量预计将突破600太瓦时（TWh），这一数值已超过加拿大全国一年的总耗电量。²同时，为满足AI系统高算力运行下的降温需求，每年耗水量或超30亿立方米，这个规模相当于挪威、瑞典等国家的年度总用水量。³除资源消耗外，AI数据中心碳排放问题同样严峻。有预测显示，未来AI领域的碳排放量将增长11倍，占全球二氧化碳总排放量的比例可能达到3.4%。（见图一）

值得关注的是，中国AI相关资源消耗与碳排放问题同样突出。有研究表明，若将运营排放与隐含排放纳入统计范围，到2030年，中国数据中心因AI应用而产生的耗电量可能超过1000太瓦时，到2038年，中国AI的年碳排放量峰值预计可能达到6.95亿吨。⁴

这种激增的资源消耗和碳排放，不仅会加剧环境破坏，还将直接推高企业AI基础设施的运营成本，同时进一步加剧部分地区电力与水资源的供应紧张局面。若该趋势得不到有效管控，还可能导致监管部门制定专项管控措施。

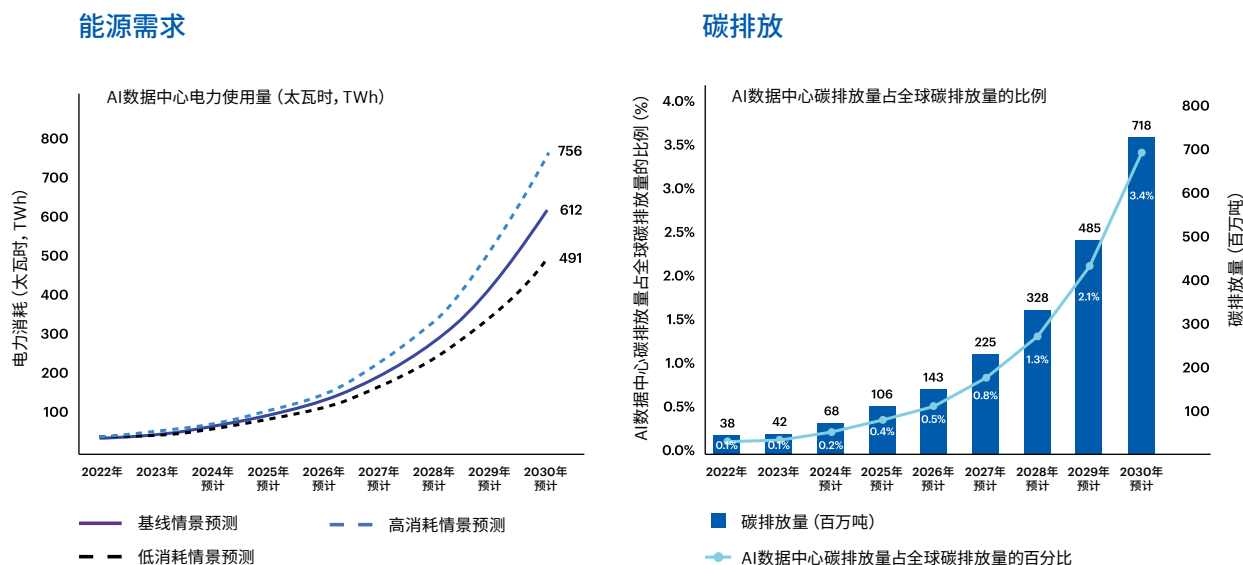
1. 《一马当先：AI规模化应用指南》，埃森哲，2025年5月6日，<https://www.accenture.cn/cn-zh/insights/data-ai/front-runners-guide-scaling-ai>。

2. 加拿大电力消耗，加拿大能源信息，2025年6月9日，[https://www.enerdata.net/estore/energy-market/canada/#:-:text=Canada%20Power%20Consumption,and%20services%20sectors%20\(27%25\)](https://www.enerdata.net/estore/energy-market/canada/#:-:text=Canada%20Power%20Consumption,and%20services%20sectors%20(27%25))。

3. 2021年年度淡水提取量，全球变化数据实验室（Global Change Data Lab），2025年6月9日，<https://ourworldindata.org/grapher/annual-freshwater-withdrawals#sources-and-processing>。

4. 《中国人工智能碳足迹预计在2030年碳达峰之后翻倍》，ScienceDirect，2025年9月3日，<https://www.sciencedirect.com/science/article/abs/pii/S0140988325007078#:~:text=Our%20projections%20indicate%20that%20the,west%2C%20creating%20new%20carbon%20hotspots>。

图一 AI数据中心的能源消耗与碳排放



来源：埃森哲商业研究院根据多个数据来源进行的估算。

可持续AI商数

解决AI环境足迹问题的首要前提是对影响进行精准量化衡量，先要算得清，才能管得好。然而，数据中心能源使用效率 (PUE) 等传统IT指标存在明显局限性，这类指标仅能反映部分资源消耗情况，无法回答核心问题，即企业为AI投入的每一份资源，最终能转化为多少实际业务价值。

为此，埃森哲提出新的衡量指标“可持续AI商数” (Sustainable AI Quotient)。作为多维度综合衡量标准，SAIQ以token (AI输出单位) 作为统一计量基准，专门用于评估AI系统将资金、能源、水资源及碳排放等投入转化成实际性能与业务价值的效率，其具体计算公式如下：

$$SAIQ = w_1 (\$/token) + w_2 (MWh/token) + w_3 (tCO_2e/token) + w_4 (m^3 Water/token)$$

企业可根据自身战略优先级，动态调整w的权重。例如，成本敏感型企业可提高财务效率 (w1, 每个token的成本) 的权重占比；强调可持续发展的企业，则可强化碳排放的影响 (w3, 每个token的碳排放量)。总之，SAIQ数值越低，表明AI系统的资源转化效率越高，可持续性表现越强。

这一衡量指标将AI可持续发展的抽象目标，转化为可量化、可管理的关键绩效指标 (KPI)。它能帮助企业管理者系统性平衡三大核心维度 (见图二)，一是成本效率，确保AI投入具备经济可行性与长期可负担性；二是能源效率，保障AI基础设施的运行韧性与能源供应安全性；三是环境效率，助力AI发展符合全球可持续性标准与监管合规要求。当SAIQ被纳入AI治理体系与绩效监控后，管理者就可以在成本控制、能源节约与减碳之间实现科学权衡，最终确保AI投资切实转化为可持续的业务价值。

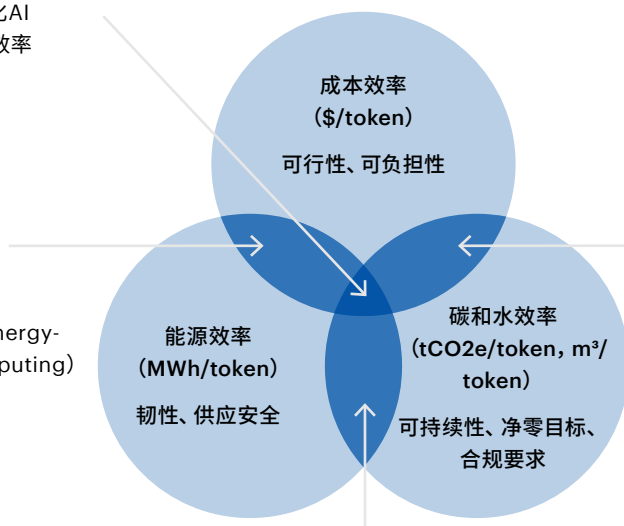
图二 优化AI效率，平衡成本、能源、碳排放和水资源影响

最佳AI性能，更低成本、能耗和排放

- 开放协作与去中心化AI
- 实时能源跟踪与碳效率

最小化能耗与成本

- 低功耗AI芯片
- AI优化架构
- 边缘AI高效计算
- 智能负载均衡
- 能源按比例计算 (energy-proportional computing)



最小化能耗与碳排放

- 轻量化模型与更智能的算法
- 循环经济中的AI
- 碳感知AI治理
- 节水型冷却系统

降低AI成本和碳足迹

- 基于使用量的AI定价模型
- 碳感知的模型部署
- AI用于减排

实现可持续AI的四大关键举措

企业要想优化SAIQ，确保AI增长具备可持续性，需从多维度协同发力。我们总结了四大关键举措，形成一套实用方案，助力AI既保持创新力，又具备可持续性。

一、优化模型设计，促进开源生态

在推动AI可持续发展的进程中，优化模型设计与促进开源生态构成了技术层面的重要抓手。从生态维度来看，中国已在全球大模型开源领域占据重要地位。据统计，中国对全球大模型开源生态的贡献度达18.7%，位列全球第二，这为模型设计优化提供了丰富的技术共享土壤与协作空间。⁵

另一方面，若要切实提升AI的资源利用效率，

还需从模型设计源头突破。具体而言，需要加快研发更高效的模型训练和推理方法，如通过模型量化、模型剪枝等轻量化方法，或者对算法进行有针对性调整，在确保模型精度不受损的前提下，大幅降低计算消耗。这种“不依赖硬件堆砌规模，要依靠设计提升效率”的思路，已得到企业实践的有效验证。

以DeepSeek为例，该公司在投入约600万美元开发DeepSeek-V3 Base模型后，又以仅29.4万美元的成本完成了R1模型的强化学习训练。即便将两部分投入合并仍远低于业内普遍动辄千万美元的竞品开发费用。⁶这一效率优势源于其优化后的混合专家架构 (Mixture-of-Experts, MoE)：在执行任务时，仅激活千亿参数中的专长子集，从而显著降低计算开销与推理延迟。同时，其独特的token分配策略发挥了关键作用——简单任务消耗更少token，复杂任务则

5. 《中国在全球大模型开源生态的贡献度达到18.7%》，第一财经，2025年9月13日，<https://m.yicai.com/news/102822638.html>。

6. 《里程碑式论文揭示深度求索 (DeepSeek) AI模型的核心奥秘》，自然，2025年9月17日，<https://www.nature.com/articles/s41586-025-09422-z>。

动态生成更多token。这种灵活机制不仅减少了对计算资源的依赖，也显著压缩了高性能大模型的整体开发成本。⁷

这些实践表明，更智能的模型设计不仅能从源头降低AI的资源消耗，还为开源生态提供了高效的技术范例，从而有效推动AI的可持续发展。⁸

二、推动数据中心脱碳，减轻环境负担

数据中心作为AI运行的核心基础设施，其能耗与碳排放是AI环境足迹的重要组成部分。推动数据中心脱碳需构建多维度协同策略。

在运营调度层面，前沿企业已探索出“动态扩展”和“智能负载均衡”技术，让服务器仅在有效算力需求时启动运行，避免闲置能耗，并优化AI工作负载的调度时间。例如，谷歌就会把部分AI任务调整至可再生能源供应充足、电价更低且碳排放更少的时段运行。⁹

在能源供给层面，一些头部科技企业正将数据中心产能规划与可持续项目深度绑定。以阿里云为例，他们通过“分布式光伏部署+清洁电力市场化交易+长期购电协议”的组合路径，系统化推进数据中心向清洁能源转型。¹⁰ 谷歌、微软、亚马逊等巨头在布局新数据中心时，会同步推进风电、光伏等可再生能源项目的落地，确保数据中心能源供给中新能源占比持续提升。¹¹ 在零碳能源这一更前沿的领域，部分企业已开始探索小型模块化核反应堆技术的应用，为未来高算力需求的云计算数据中心提供近乎零碳排放的稳定能源。¹²

同时，先进冷却技术方案成为数据中心降耗脱碳的重要助力。其中，液冷技术正逐步替代传统风冷技术，成为AI时代解决高算力芯片散热难题、降低冷却能耗的主流方向。在实践应用中，头部科技企业探

索已验证了液冷技术的可持续价值。例如，微软试验的“双相浸没式冷却”方案，¹³ 以及谷歌在部分数据中心采用再生水进行冷却的实践，¹⁴ 均证明AI基础设施在维持高性能运算的同时，可有效减少对传统机械制冷设备的依赖及淡水资源的消耗，实现算力性能与绿色低碳的协同发展。

三、战略性部署AI，用AI实现脱碳

AI的部署需依托具体应用场景制定针对性策略，核心目标是实现效益最大化和资源浪费最小化。当前部分企业存在明显的误区，即盲目追求大模型应用，无论任务复杂度高低均采用大模型处理，导致计算资源、能源与经济成本的无效消耗。

战略性部署AI的核心逻辑在于适配性，让适当的模型处理适当的任务，避免资源错配。具体可通过三项策略落地：其一，借助检索增强生成（RAG）技术优化数据调用模式，仅在核心需求环节调取相关数据，减少冗余计算；其二，针对常规性、标准化任务，优先选用轻量化领域专用模型，或采用“混合AI+规则驱动”的复合方案，以更低资源消耗满足需求；其三，仅当面临高复杂度、高价值的核心问题（如多维度决策优化、复杂场景预测等）时，才动用高性能AI系统，确保资源向高价值场景集中。

除模型适配外，机制创新也是推动AI可持续部署的重要抓手。例如，企业可推动AI服务定价模式从传统“固定费率”转向“按实际使用量计费”，通过经济杠杆激励用户合理规划使用需求，倒逼资源利用效率提升。

更为关键的是，需充分释放AI自身对脱碳的赋能价值，利用AI技术优化能源消耗、减少碳排放，形成“以AI促脱碳”的正向循环。当前该领域仍存在

7. 《DeepSeek-R1通过强化学习激励大型语言模型进行推理》，自然，2025年9月17日，<https://www.nature.com/articles/s41586-025-09422-z>。

8. 《效应：通过算法效率改写AI经济学》，澳大利亚机器学习研究所（AIML），2025年3月26日，https://medium.com/@aiml_58187/the-deepseek-effect-rewriting-ai-economics-through-algorithmic-efficiency-part-1-46cf9b2e9930。

9. 《数据中心碳感知计算》，arXiv，2021年6月11日，<https://arxiv.org/abs/2106.11750>。

10. 《环境、社会和治理报告（2025）》，阿里巴巴，<https://www.alibabagroup.com/zh-HK/esg>。

11. 《可持续技术：数据中心如何通过创新应对气候变化》，Datacenters.com，2024年8月13日，<https://www.datacenters.com/news/sustainable-tech-how-data-centers-are-tackling-climate-change-through-innovation>。

12. 《谷歌押注核能，与Kairos Power签订协议，部署500兆瓦小型模块化反应堆舰队为数据中心供电》，《电力杂志》，2024年10月14日，<https://www.powermag.com/google-bets-big-on-nuclear-inks-deal-with-kairos-power-for-500-mw-smr-fleet-to-power-data-centers/>。

13. 《为了冷却数据中心服务器，微软转向使用沸腾液体》，微软，2021年4月6日，<https://news.microsoft.com/source/features/innovation/datacenter-liquid-cooling>。

14. 《谷歌2024年环境报告》，谷歌，2024年7月，<https://sustainability.google/reports/google-2024-environmental-report/>。

巨大潜力。据统计，仅有14%的企业运用AI降低碳排放，有待挖掘的价值空间广阔。¹⁵ 在实践中，AI的脱碳赋能已展现明确成效，例如通过AI算法优化建筑暖通空调（HVAC）系统的运行参数（如动态调节温度、风量等），可降低25%的能源消耗，同时保障环境舒适度，实现节能与体验的双赢。¹⁶

总而言之，AI的部署不应是缺乏战略规划的盲目推广，而需聚焦关键场景，锚定核心需求，在追求更高投资回报的同时，通过科学部署与技术赋能推动环境可持续性，最终实现高效部署AI与推进脱碳目标的协同。

四、“代码化治理”，从源头管控风险

可持续AI离不开对AI生命周期的有力治理，核心在于将可持续需求从AI开发初期便嵌入算法设计和工作流程。具体而言，就是要把环境因素融入自动化系统。其一，将AI碳强度（如单位token碳排放）和业务KPI同步纳入追踪范畴，实时监控AI运行中的能源消耗、水资源占用等核心数据；其二，采用“策略即代码”框架将可持续性阈值转化为可执行的代码规则，动态把控资源使用边界，一旦超出阈值便触发预警或自动调控。

例如，在物流场景中，通过AI实现预测性物流规划，企业能够有效预测和管理运输环节的碳排放。京东物流整合140种交通运输载具和2000多种供应链碳排放因子库，实现了99.5%碳足迹核算的准确性，通过采用节能诊断算法，“北斗+融合定位”等技术，实现“监测-报告-核查-跟踪”（MRV-T）全链条管理闭环，攻克碳排放动态监测盲区，有效降低物流运输路径的碳排放偏差，为AI在垂直领域的可持续治理提供参考。¹⁷

对企业而言，可持续AI绝非临时补救或一次性补偿可实现的目标，而需将其以代码形式嵌入AI开发、部署、管理的全流程。完善的治理机制，既能帮助企

业实时识别AI运营中的低效问题，避免碳排放突然飙升，又能提前适配全球各地的可持续监管要求，更能在负责任的前提下，安全实现AI规模化应用，为可持续AI的长期落地夯实制度与技术基础。

从资源效率视角重新审视AI价值，再用SAIQ等科学指标进行衡量，企业的可持续发展将从模糊的战略愿景转化为AI战略中可落地、可追踪、可优化的具体成果。那些主动优化AI能源消耗和碳排放的企业，更有可能实现更低的运营成本、更稳定的能源供应，还能提前适应未来的气候法规。而那些忽视AI环境足迹的企业，很快可能面临成本上升、合规等难题，甚至因为数字创新的可持续性不过关，丧失市场信任。

这种增长与可持续的平衡，对中国而言更具特殊战略意义。中国已明确“2030年前碳达峰、2060年前碳中和”的“双碳”目标，为可持续发展划定了清晰的时间线。¹⁸ 下一代AI解决方案必须将可持续性作为核心设计原则。若中国企业能优化模型设计、加速绿色数据中心建设、完善AI治理体系，不仅能推动自身实现低碳转型，更有望在全球可持续科技竞争中抢占先机，为全球AI产业可持续发展提供中国方案。■

桑杰·波德尔 (Sanjay Podder)

埃森哲董事总经理、全球技术可持续创新负责人

沙拉布·库马尔·辛格 (Shalabh Kumar Singh)

埃森哲商业研究院高级研究总监

马修·吉朱 (Mathew Giju)

埃森哲商业研究院研究总监

张逊

埃森哲大中华区战略与咨询事业部董事总经理、可持续发展业务主管

玛尼莎·达什 (Manisha Dash) 和于雅对本文亦有贡献

业务垂询: contactus@accenture.com

15. 《净零目标》，埃森哲，2024年11月10日，<https://www.accenture.com/us-en/insights/sustainability/destination-net-zero>。

16. 《Dollar Tree借助BrainBox AI实现大幅节能减排》，BrainBox AI，2025年6月9日，<https://brainboxai.com/en/case-studies/dollar-tree-unlocks-major-energy-and-emissions-savings-with-brainbox-ai#:~:text=This%20impressive%20reduction%2C%20driven%20by,orders%20raised%20at%20these%20locations.>

17. 《入选国家示范清单！京东物流以MRV-T技术推动物流行业绿色低碳转型》，京东物流，2025年5月9日，https://mp.weixin.qq.com/s/_hCoQJGEIx4Wda6UjTEqVg。

18. 《中国有望在2060年目标之前实现碳中和》，哈佛商学院，2025年6月20日，<https://www.hbs.edu/bigs/china-poised-to-meet-carbon-neutrality-goal-before-2060>。